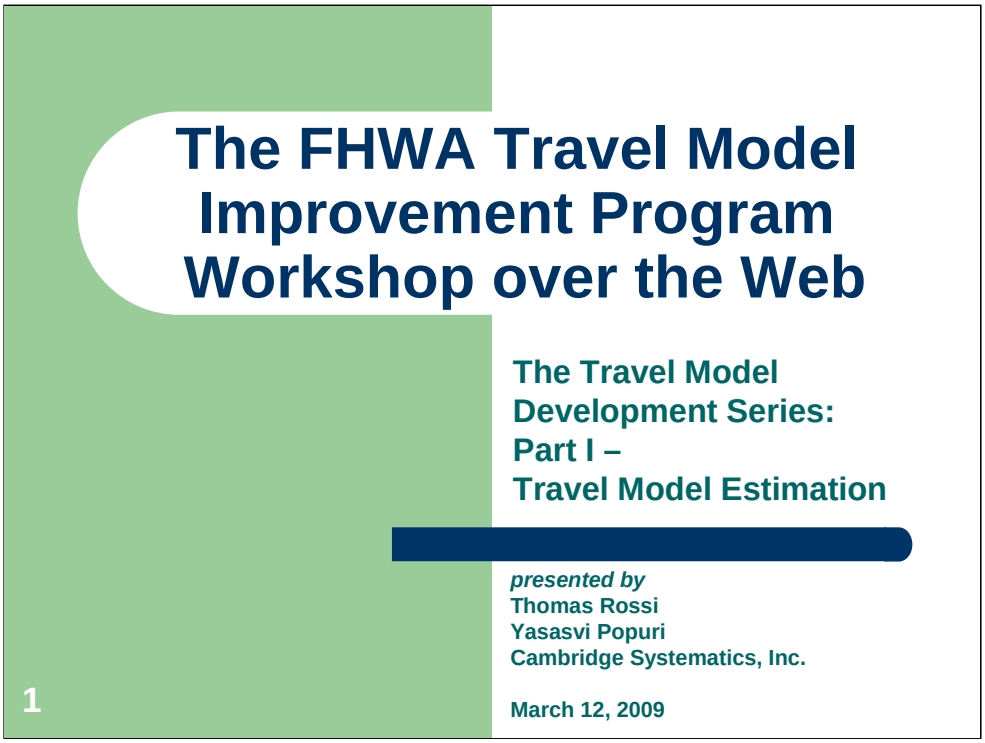
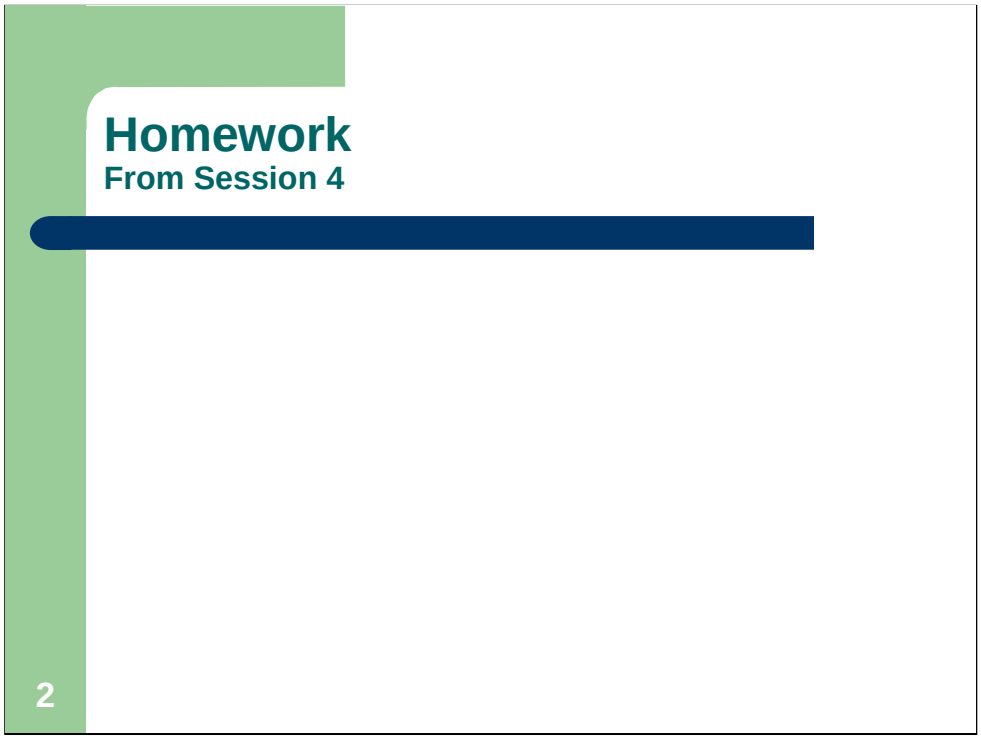


Key Message: Purpose of the Webinar Series

Details:

Welcome to the FHWA TMIP Workshop over the Web. This workshop is targeted at Transportation modelers who have a low to moderate level of familiarity with the estimation and validation of travel models.

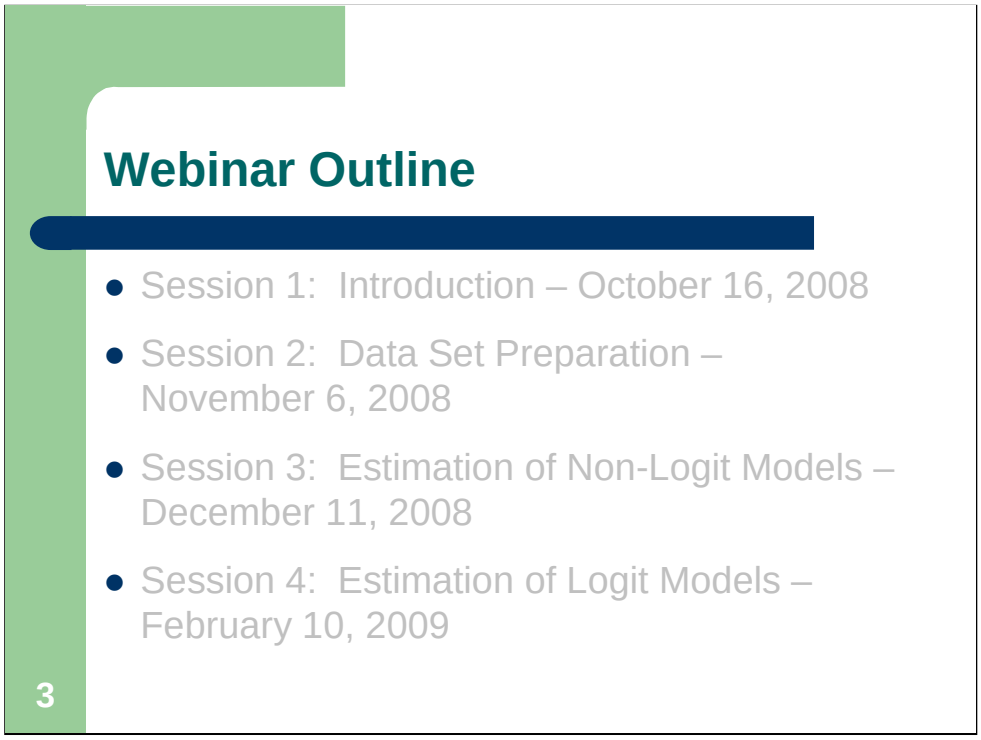
This series of webinars will introduce the development of model estimation data sets, the structures of the various model components, and the procedures for estimating models. The workshop will include lectures, discussion, and "homework," that participants will be expected to complete between sessions.



Key Message: Homework Discussion

Details:

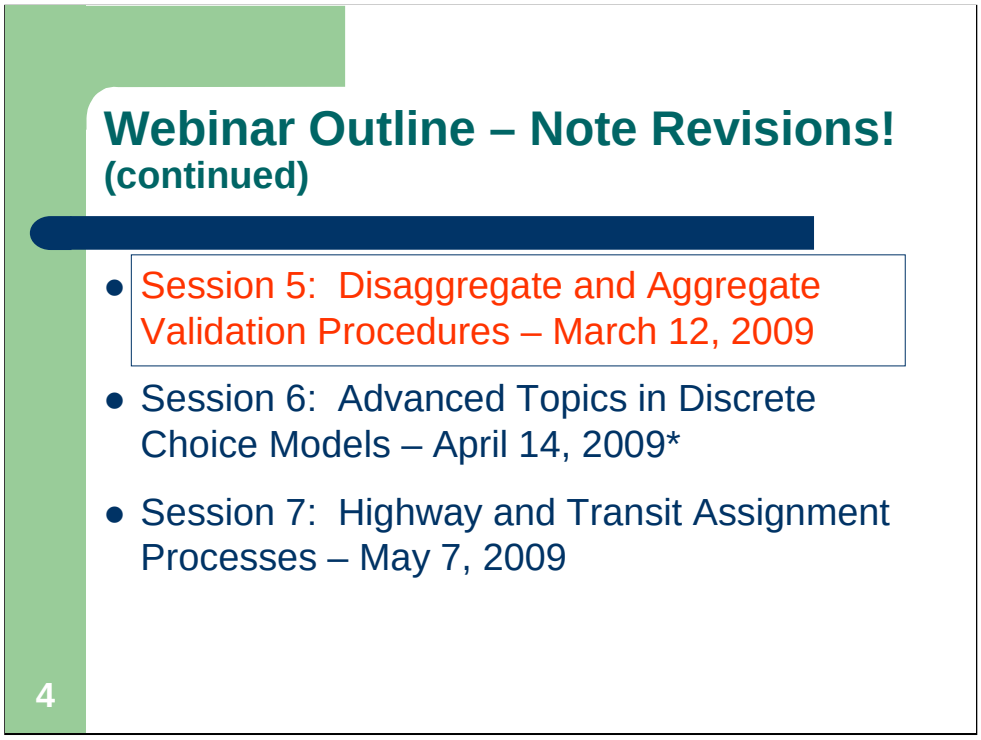
Please refer to the homework solutions posted at the website.



Key Message: Past Sessions

Details:

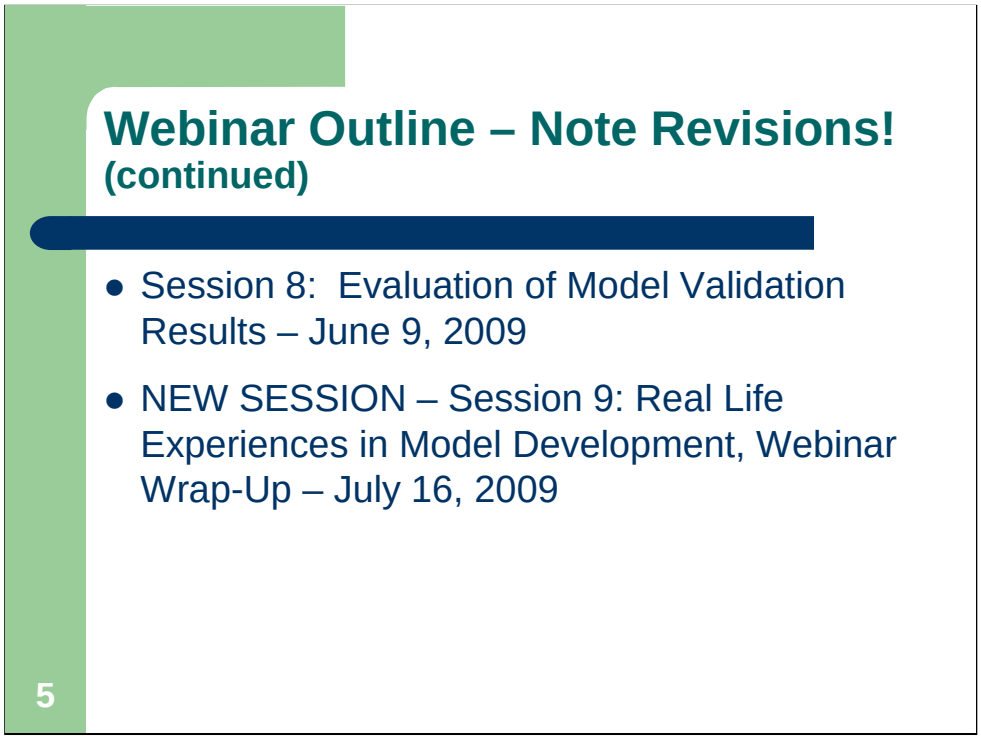
This is the webinar outline. These are the sessions that have already occurred.



Key Message: Upcoming Sessions

Details:

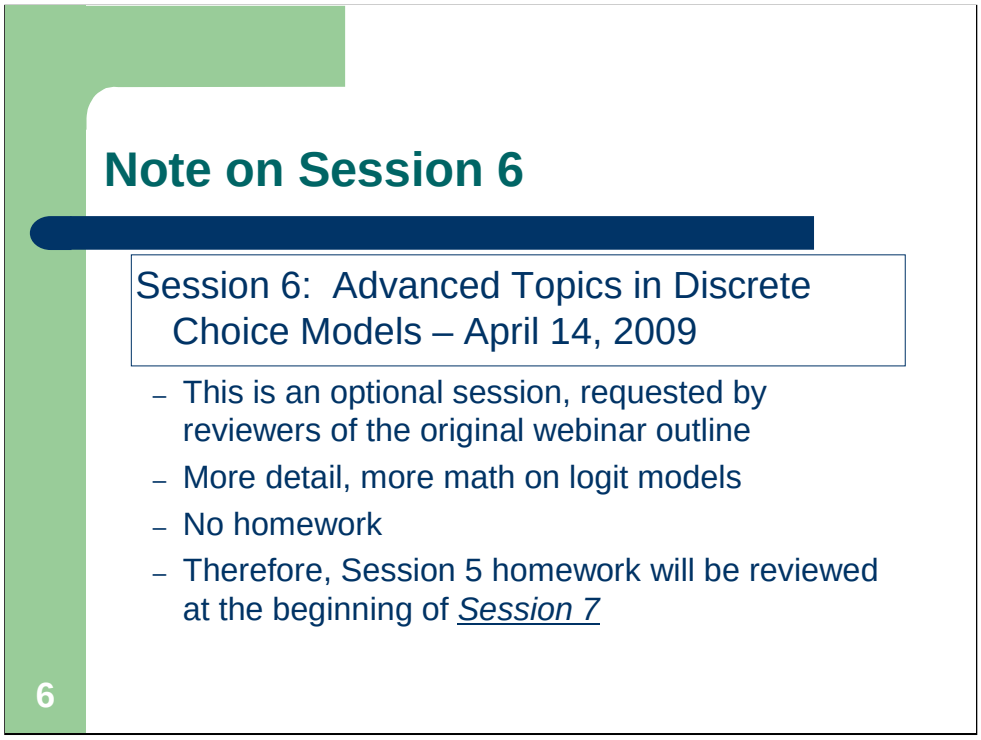
After today's Session 5, Session 6 will be conducted on April 14, and Session 7 on May 7.



Key Message: Upcoming Sessions

Details:

Session 8 will be conducted on June 9 and a new Session 9, on real life experiences in model development and the webinar wrap-up, will be held on July 16.



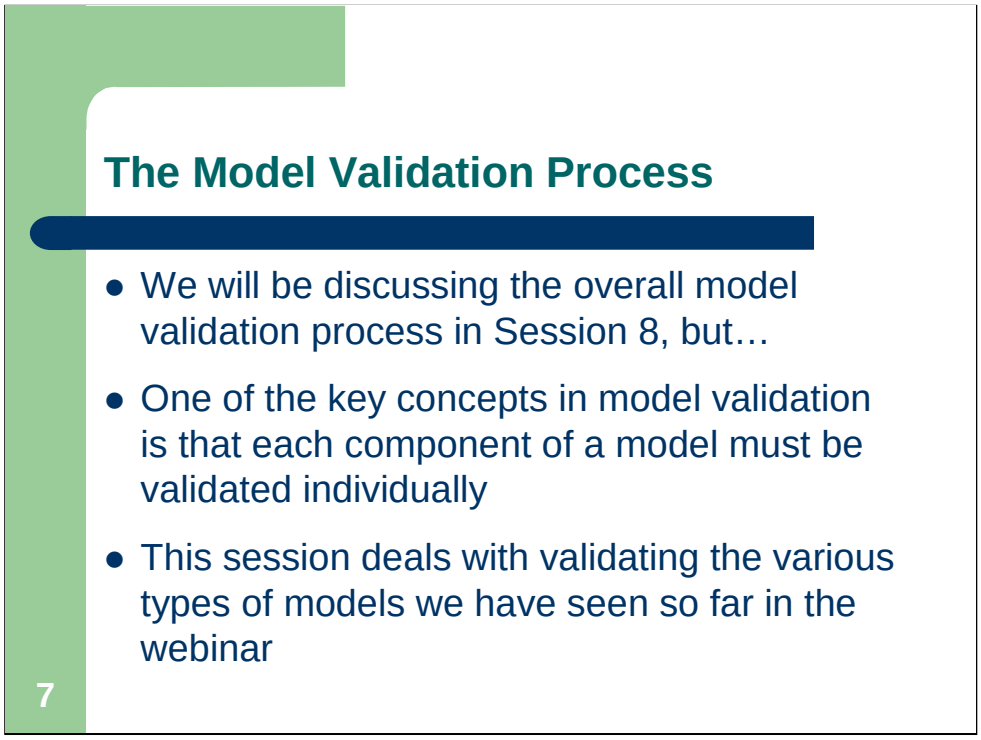
Key Message: Session 6

Details:

A note on next month's session. This was requested by reviewers of the original outline for this webinar for people who want to get into more detail on logit models. Topics will include:

- Modeling disaggregate individuals
- More on generic vs. alternative-specific variables
- Interpreting model estimation results
- Examples of advanced variable specification
- Likelihood functions
- Application programming for logit models

Because this is an optional session, there is no homework associated with it. Therefore, the homework for this session will be reviewed at the beginning of Session 7 in May.



Key Message: The Model Validation Process

Details:

Later, in Session 8, we will discuss the overall process of validating the entire model system. But one of the key concepts in the validation process is that each and every component of a model must be validated individually. So this session will deal with the validation of model components that we have seen so far in the webinar. These include:

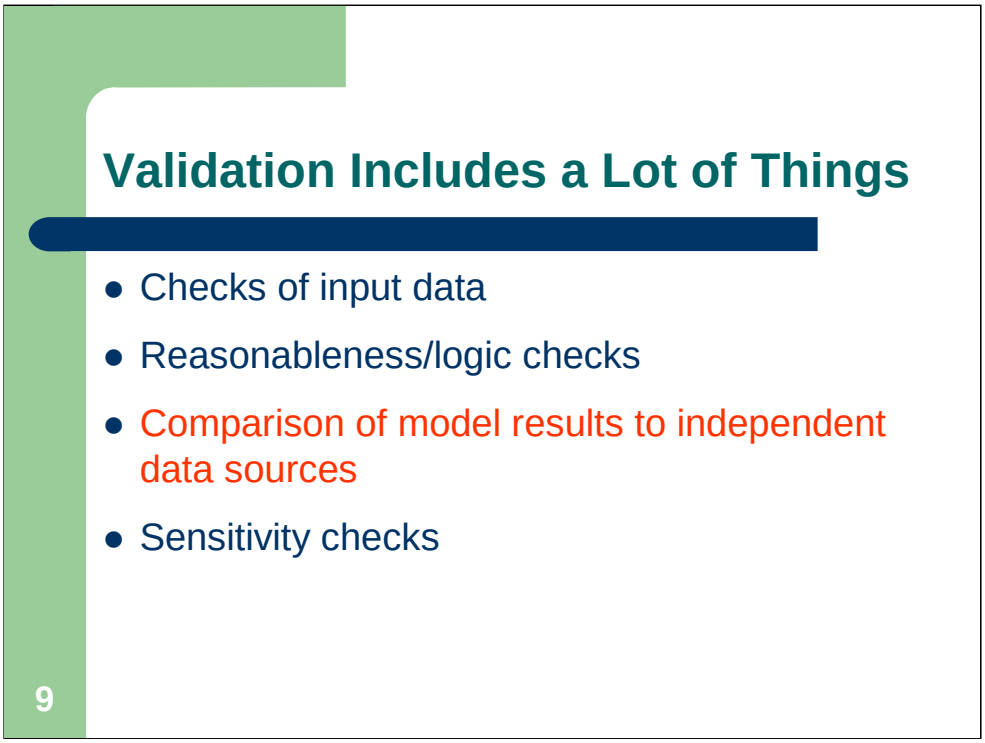
- Trip generation models
- Trip distribution models (e.g. gravity)
- Logit models, including mode choice

It is important to note that this session will likely leave you "wanting more." We cannot cover the entire model validation process, with lots of examples, in less than two hours. For example, we do not have time to cover validation standards or typical values for tests. We do hope, though, that it whets your appetite for the complete documentation of model validation that FHWA is currently working on, and that will include case studies and outreach over the web.

The New FHWA Model Validation Manual Is Coming!

- Later this year!
- Rejected titles for the new manual:
  - "Son of Model Validation and Reasonableness Checking Manual"
  - "Validation Redux!"
  - "Validation II – The Sequel!"
  - "Validation Wars Episode 5 – The Modeler Strikes Back"

8



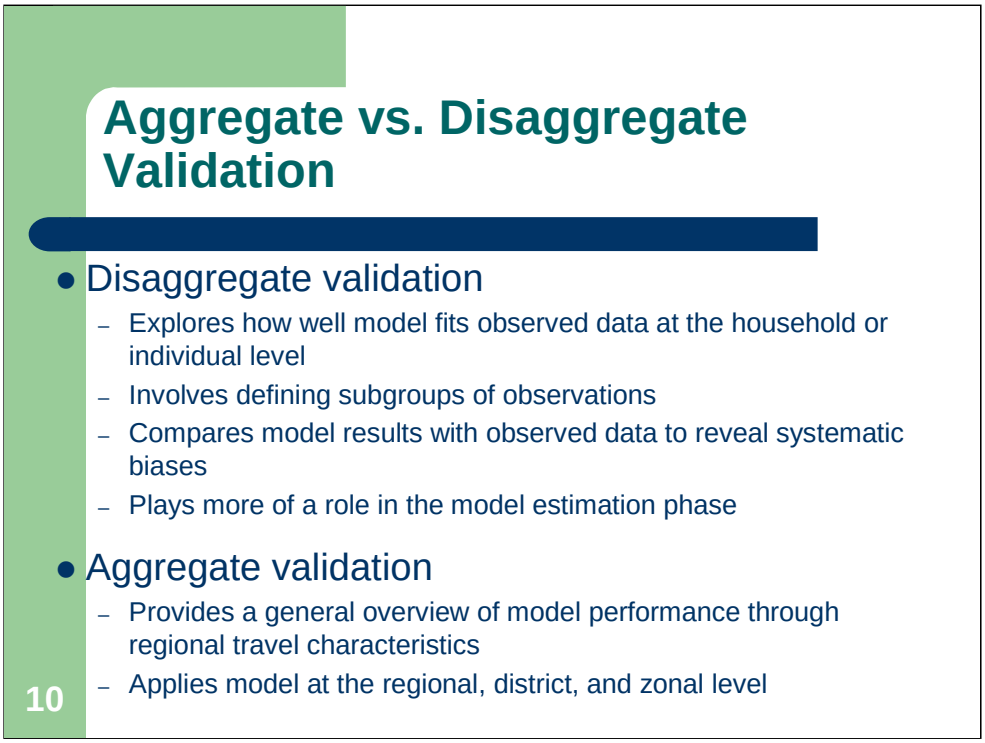
Key Message: The Model Validation Process

Details:

Model validation includes a lot of things. This includes checking the input data for accuracy, which we won't be covering today. We will be covering some, but not all possible reasonableness or logic checks. Some of those checks include things we discussed during the discussion of model estimation in the last two sessions, such as checking error levels, statistical significance, and estimates against previous experience.

A key component of model validation is comparing results to independent data. For example, if you had a survey that was large enough, you could divide it into two independent parts, one for estimation and one for validation. Most of the time, we're lucky to have enough data for estimation, though. There are other independent data sources we use, including traffic counts, transit ridership, census data, etc.

We'll also discuss sensitivity checks toward the end of this session.

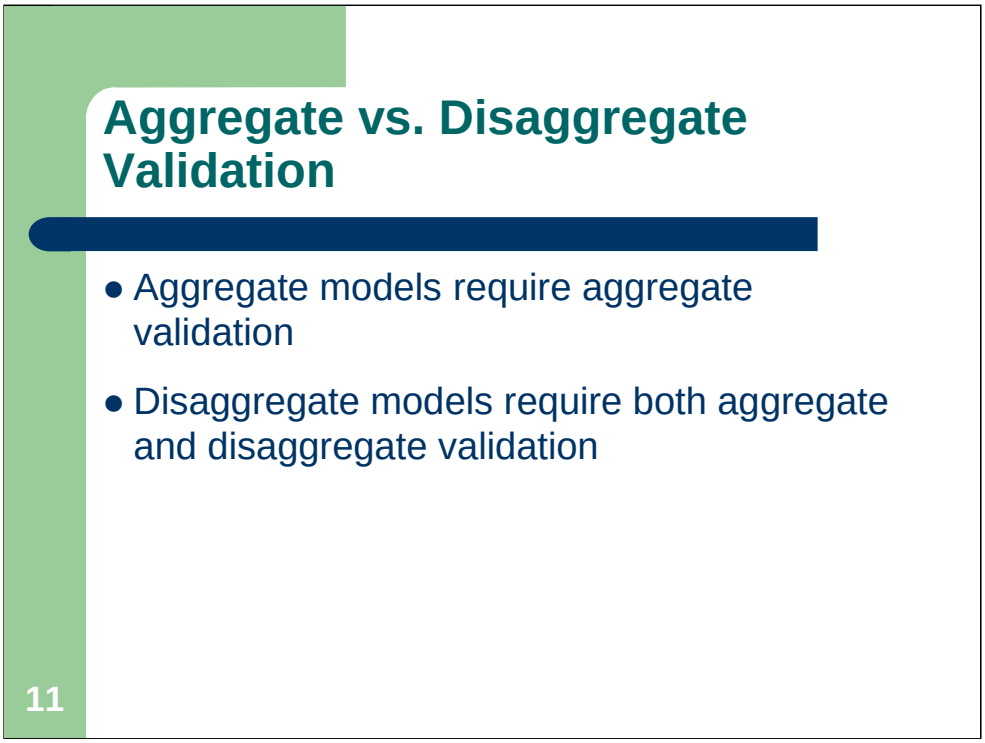


Key Message: Aggregate vs, Disaggregate Validation

Details:

Disaggregate validation focuses more on comparing the predicted and observed behavior at the *individual or household* level. An example would be to see if the predicted mode share of individuals with 1 vehicle in their household compares closely to the observed data.

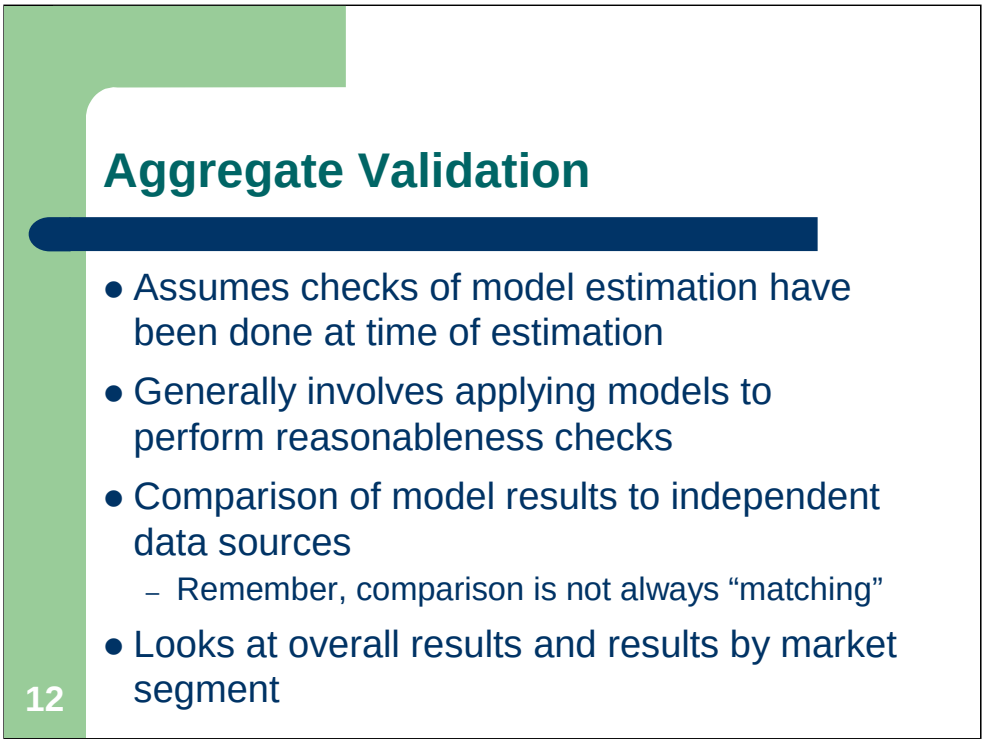
Aggregate validation focuses on regional validation. For example, comparing the predicted trip length distributions to the observed trip length distributions from a household survey could be categorized as disaggregate validation.



Key Message: Aggregate vs. Disaggregate Validation

Details:

Please see the previous slide.



Key Message: Aggregate Validation

Details:

Aggregate validation processes assume that the model estimation checks have already been done at the time of the estimation. Aggregate validation generally involves applying the model and then making reasonableness checks using observed data. Going back to the example of trip distribution, we would use aggregate validation in the context of comparing, say, trip lengths by purpose and car-ownership level to observed trip lengths by those same categories.

Cross-Classification Model – Trip Productions
Review

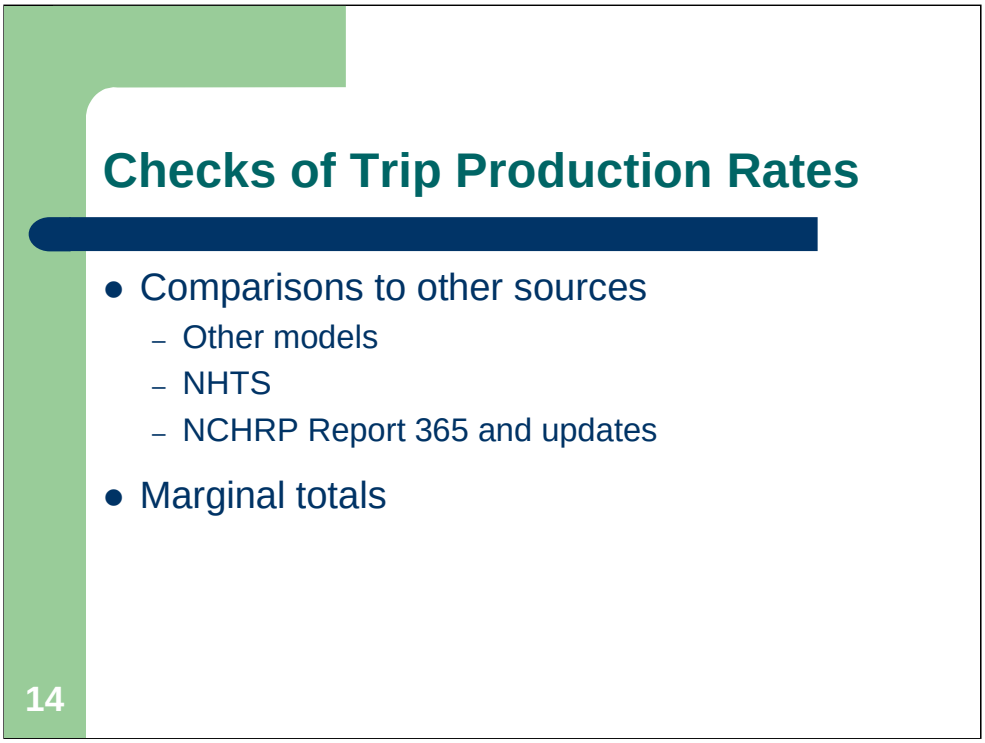
		Independent Variable #1				
		Value 1	Value 2	…	Value n	Total
Independent Variable #2	Value 1	Dep var value	Dep var value		Dep var value	
	Value 2	Dep var value	Dep var value		Dep var value	
	…					
	Value n	Dep var value	Dep var value		Dep var value	
	Total					

13

Key Contents: Cross-Classification Models - Review

Details:

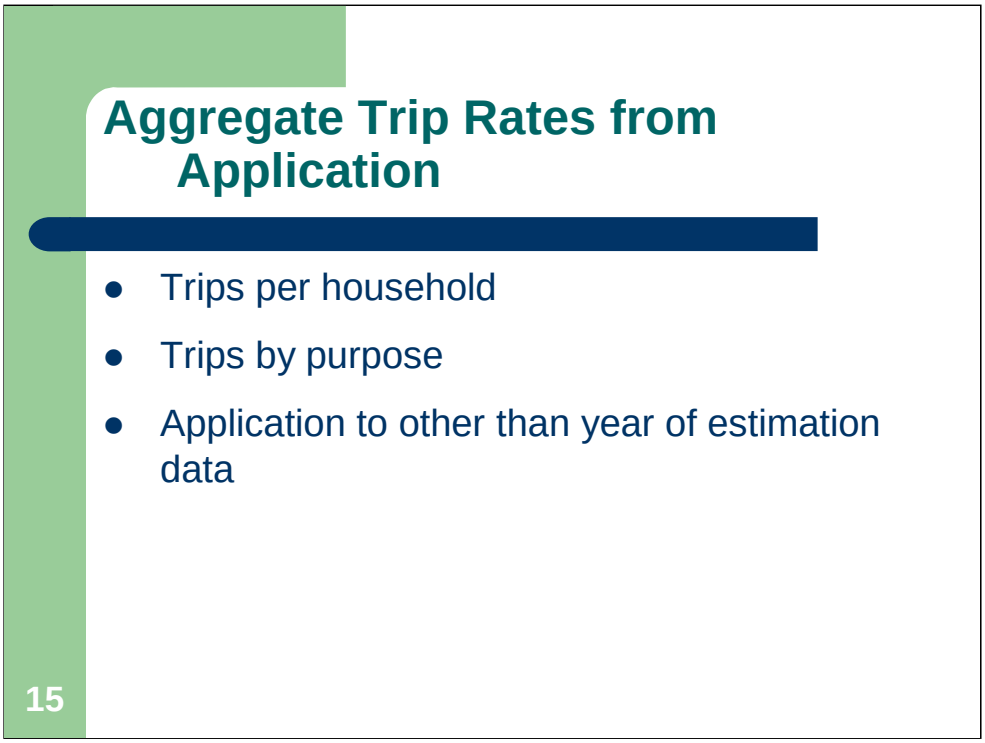
As you may recall from Session 3, cross-classification is the most commonly used technique for trip production models. Cross-classification involves identifying independent variables that will be used to classify households, identifying the number of households in each category, determining how many trips of each purpose made by the households of each type, and computing the trip rate for each cell.



Key Contents: Cross-Classification Models – Data Sources

Details:

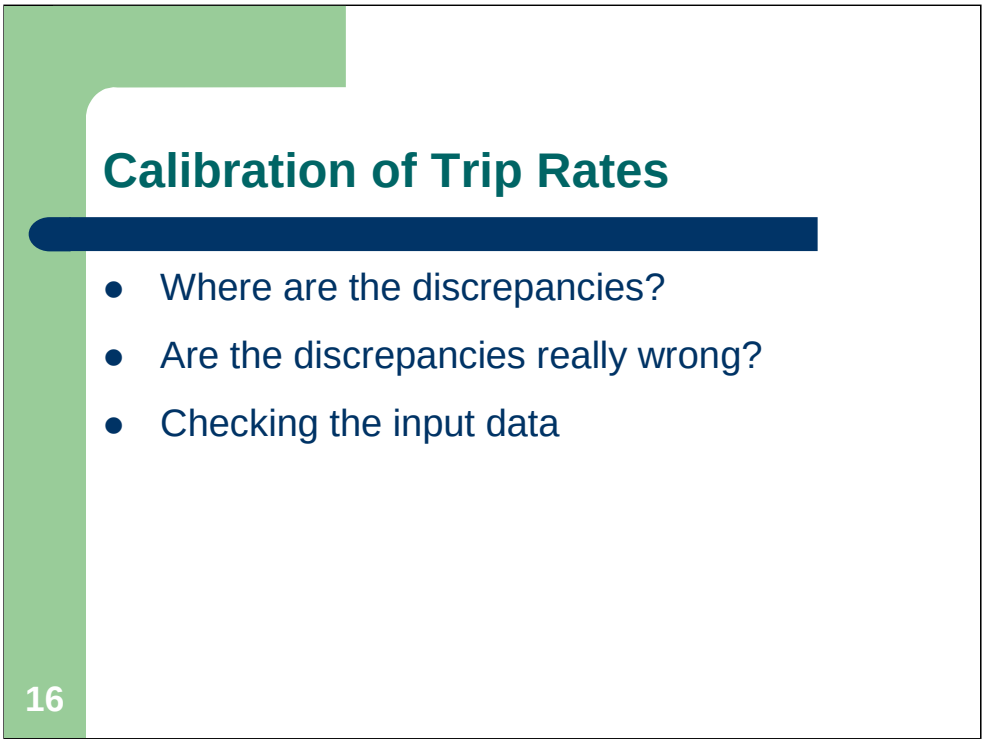
So, where can we find data to compare and validate cross-classification rates? There are three potential sources: first, the modeler could look at trip rates for similar areas elsewhere in the nation; second, one could look at the rates implied by the NHTS data; third, one could also compare the trip rates with those in NCHRP 365.



Key Contents: Cross-Classification Models – What to Look For?

Details:

Which exact trip statistics should we be looking at for the aggregate trip rate validation. Typically, the modelers look at trips per household by household type and purpose. Another good check will be to apply these rates to data for another application year and look at the changes in total predicted trips.



Key Contents: Calibration of Trip Rates.

Details:

Once the rates have been examined and compared to other sources, they could be adjusted to account for unexplainable differences. But the real test occurs when the model is applied.

For example, let's say a model applied to the model's socioeconomic input database comes up with an average HBW trip rate of 2.5 trips/household, but the rate from the survey is 2.0.

Where are the discrepancies? The implied rates from the application foe each cell must be the same (unless there is a programming error—which is something to look out for!). So the distribution of households to cells could be significantly different between the survey and the model input data. Is that necessarily incorrect? If the survey data is for the same year as the model base year, then it probably is, since the survey data should be expanded to match the total population. But if the application is for a different year (forecast or backcast), then the distribution could change. It is worth checking to see if the change is correct (perhaps certain segments are growing faster than others) or if the data for one or both years is erroneous. If the data are correct, then the modeler needs to determine if the change in trip rate is reasonable. It may well be.

Regression Model – Attractions Review

$Y = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_n X_n$

where:

Y = Dependent variable

B_i = Estimated coefficients

X_i = Independent variables

The maximum likelihood estimators for coefficients are based on method of least squares

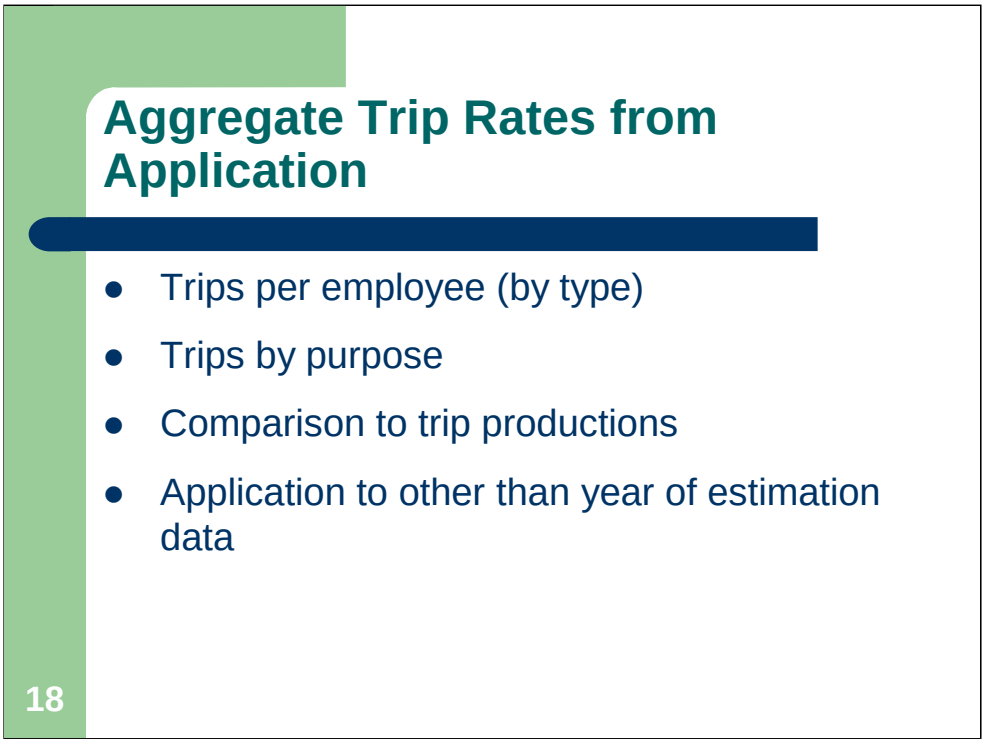
17

Key Message: Regression Model - review

Details:

You may recall the regression model, which we covered in Session 3 and is typically used for trip attractions.

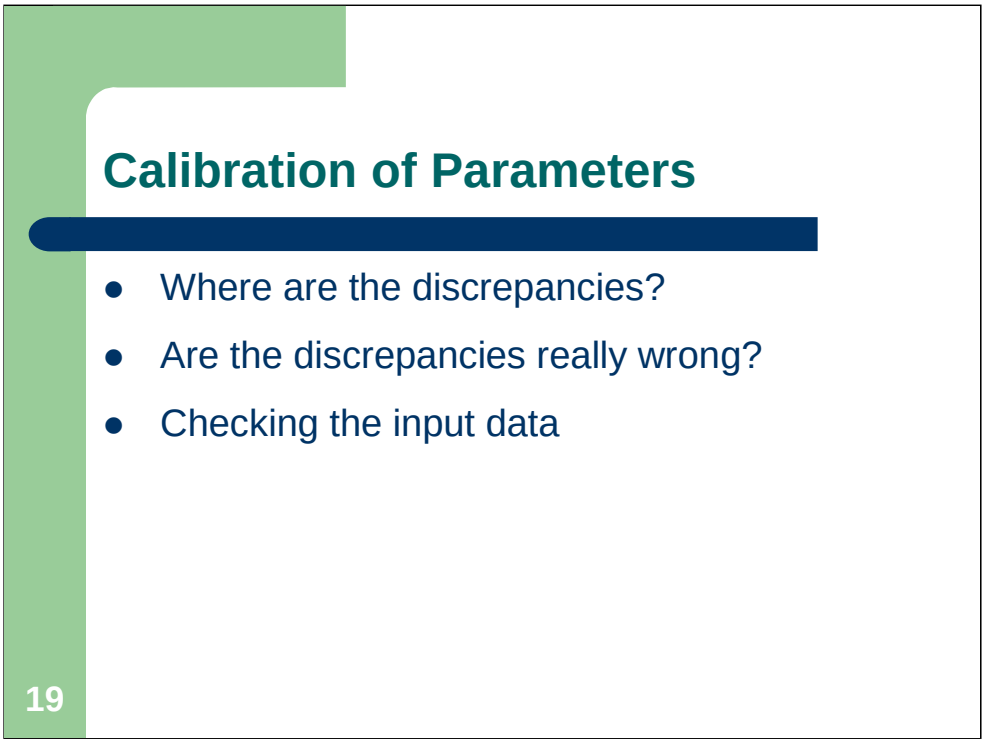
The number of trip attractions is a linear function of one or more variables describing the level of activity in a zone, such as employment by type or the number of households.



Key Message: Trip Attraction Rates from Application

Details:

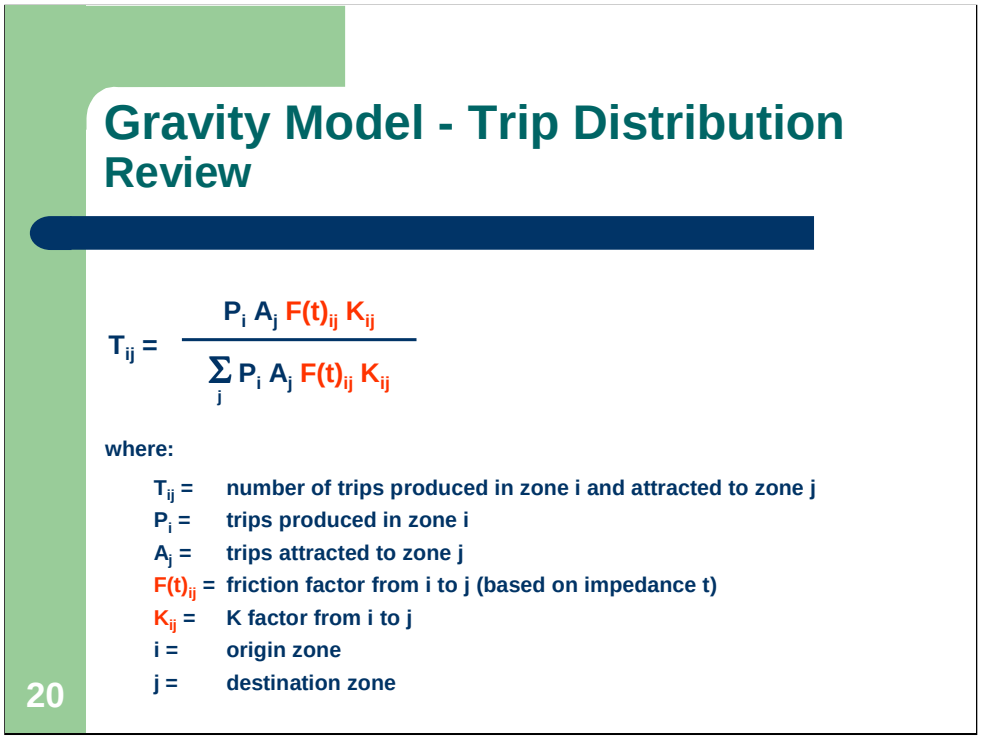
The trip attraction model generates estimates of trips per employee by employment type and purpose. Because the trip attraction rates are based on employment data, they use a fundamentally different set of data. After applying the trip attraction rates, it is a good idea to compare the number of trip ends obtained to those obtained by applying the trip production rates. Also, as with the trip production rates, it is always a good idea to apply the trip attraction rates to data from some other year and look at the changes in total predicted trips.



Key Message: Trip Rates – Other Checks

Details:

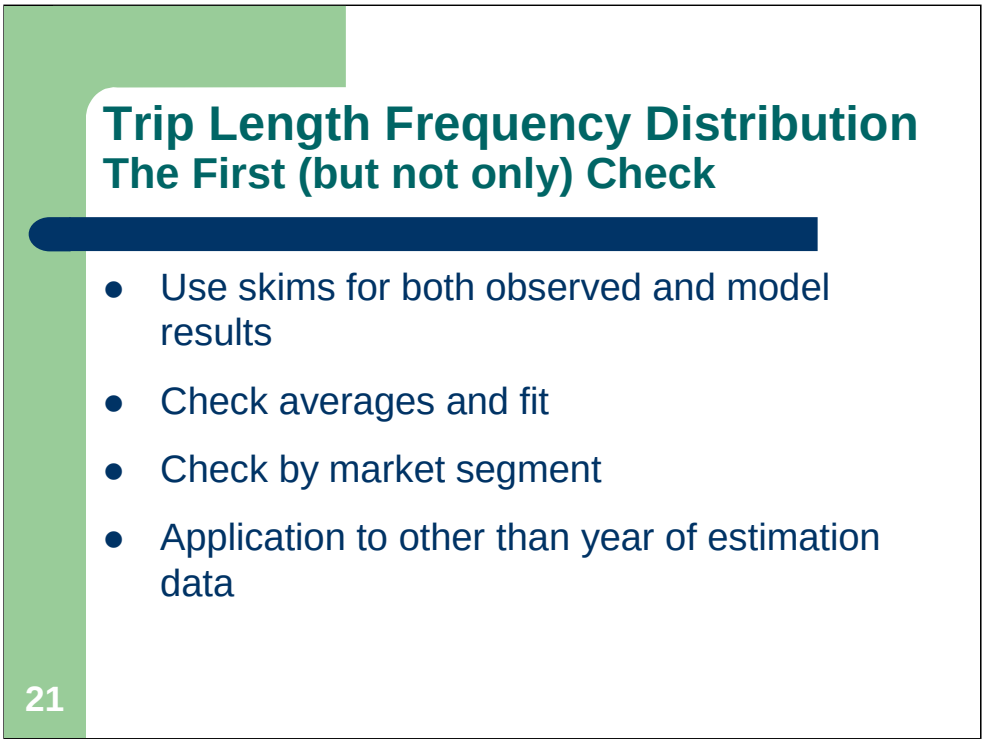
Whether we are checking the trip production rates or the trip attraction rates, it is important to understand the sources of the discrepancies. For example, if households of a given cross-class category appear to have really high trip rates, we need to question if this is indeed a true pattern or if this has to do with the lack of data records, and therefore biased upwards.



Key Contents: Gravity Models for Trip Distribution

Details:

The gravity trip distribution model was also covered in Session 3. The estimated parameters are the friction factors, or the parameters of the distribution (e.g. gamma function) chosen for the friction factors. We will discuss K-factors shortly.



Key Contents: Gravity Models for Trip Distribution

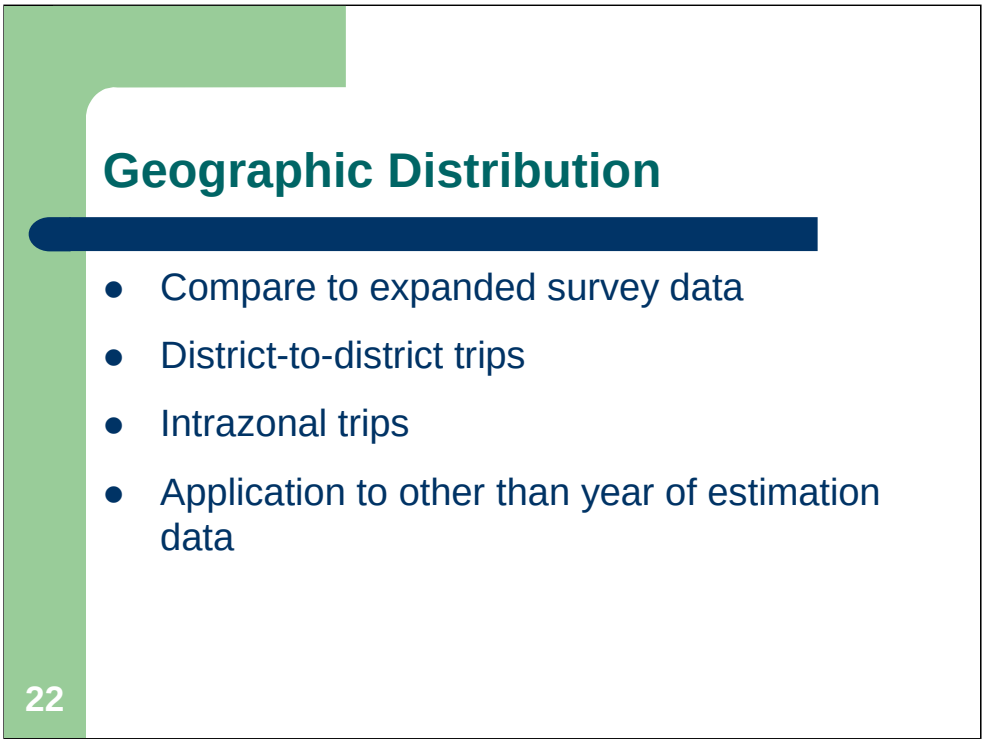
Details:

Some people make the mistake of doing only trip length comparisons to validate the trip distribution model. It is important, but not the only check. If there is a local household survey data set available, it is the best validation data source although if it was also used for model estimation, it is not an independent source.

The comparison is done using the trip length frequency distribution, usually by minute and is done separately for each trip purpose. It is also easy to check the average trip length by purpose. One way of checking the fit of the trip length frequencies is the coincidence ratio, where both frequencies are plotted and the area under both curves is divided by the area under either curve. It is important to apply the network skims to the survey data based on the origin/destination of each trip record rather than using reported travel times from the survey, which are often inaccurate and always “lumpy.”

Checks can be done not just for all trips of a purpose, but by market segment (geographic, demographic, etc.)

Applying the model for more than one year is a good idea. Do the trip lengths change significantly? Do these changes make sense? For example, is there more growth in outer areas of the region in one scenario, making trip lengths longer?



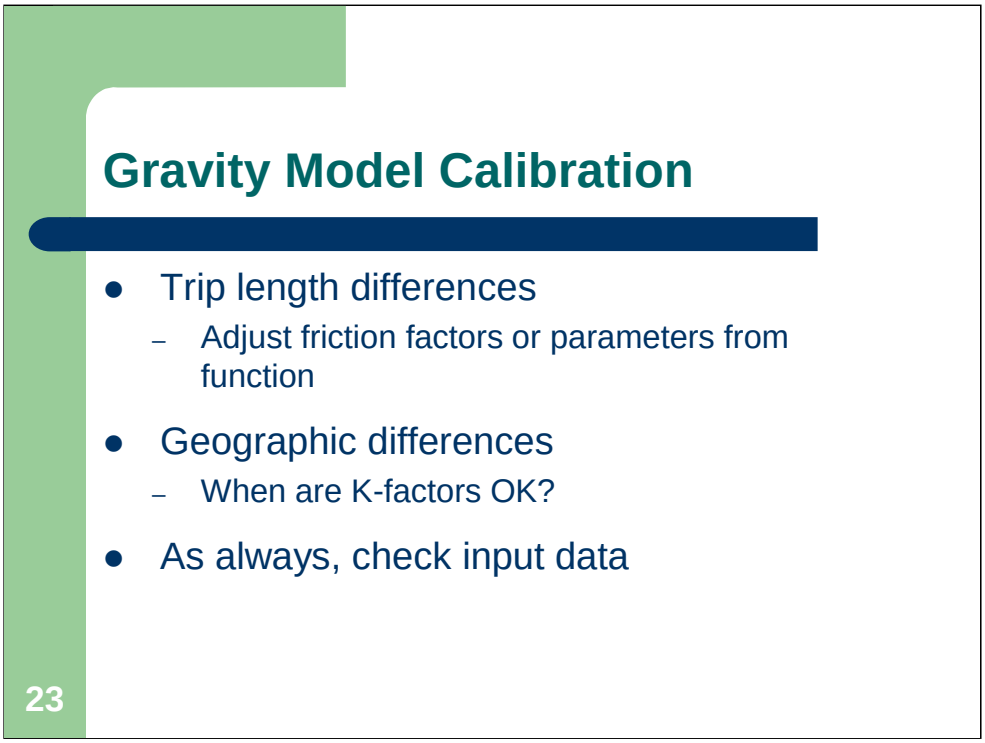
Key Contents: Gravity Models for Trip Distribution – Key Statistics

Details:

The validity of Gravity model outputs can be checked using various observed statistics.

First, the study area could be divided into various districts based on regional importance. Then the trip table could be aggregated from the TAZ-level to the district-level. This more aggregate trip table can now be compared with the expanded district-district table from household surveys.

Second, one could also look at the proportion of trips that are happening within a TAZ. This proportion will naturally vary by purpose. For instance, the work trips tend to be longer and therefore, a smaller proportion of work trips tend to be intrazonal. Comparing the modeled fraction of intrazonal trips by purpose to the observed data from the survey can yield useful insights about the model performance.



Key Contents: Gravity Model Calibration

Details:

So, how do we adjust the gravity model parameters?

Well, the modeled average trip lengths and the observed average trip lengths can be compared for each purpose. Then, the friction factors can be adjusted to be more or less onerous to match the observed trip lengths.

Along the same line, it is possible to compare the modeled and observed district-district tables and generate K-factors that align the modeled patterns with the observed patterns. Of course, one should not blindly create K-factors and use them in the model. Even before any K-factors are introduced, it is important to check the input data and understand the reason for the differences in the modeled and observed patterns

Logit Models
Review

Probability function:

$$P(1) = \frac{\exp(v_1)}{\exp(v_1) + \exp(v_2) + \cdots + \exp(v_n)}$$

Used for:

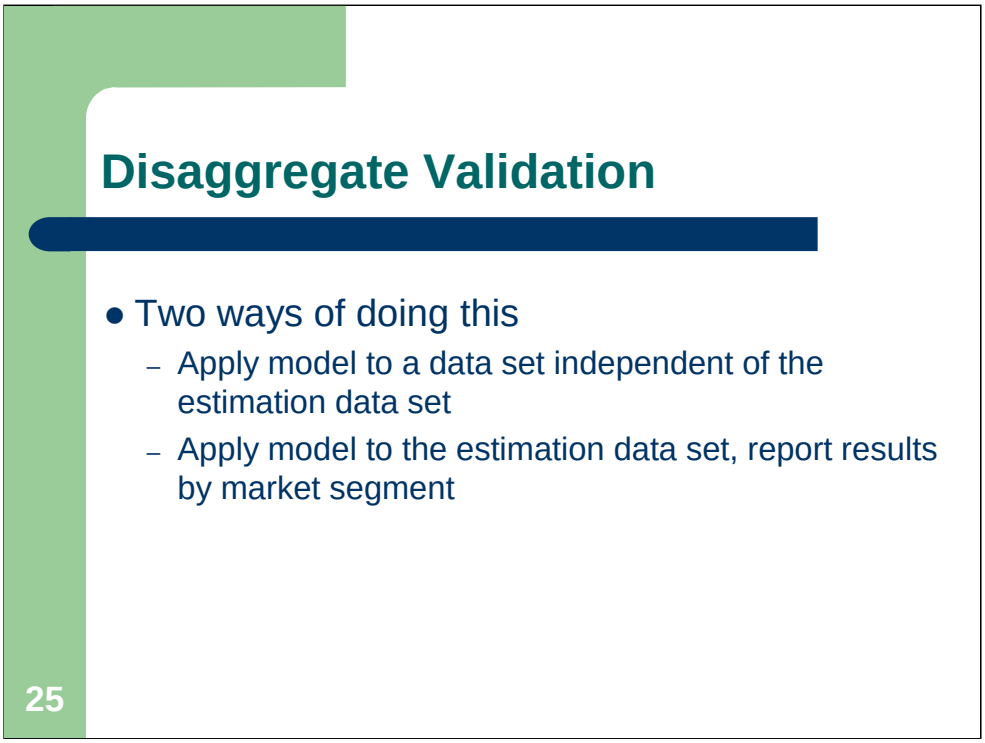
- Mode choice
- Vehicle availability
- Destination choice
- And others…

24

Key Message: Introducing Validation of Logit Models

Details:

Next we will discuss the validation of logit models. As you recall from the last session, logit models are used to model any choice with discrete alternatives (modes, number of vehicles available, destination zones, etc.). While mode choice is by far the most common application for logit models in travel modeling, they can be used in many other types of models.

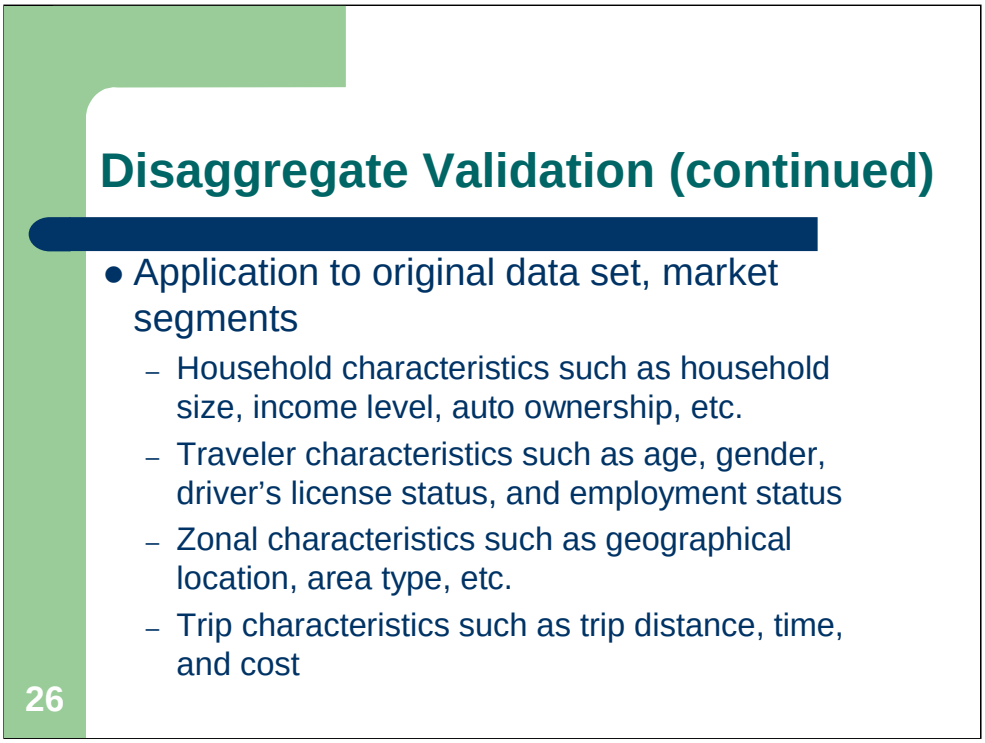


Key Contents: Disaggregate Validation

Details:

Disaggregate validation can be done in one of two ways. First, we could apply the model to a completely new data set and compare the predicted shares with the observed shares in the new data set.

If data are limited, we could go through the same exercise on the estimation data set itself, and report results by market segment to see where the model is performing poorly.



Key Contents: Disaggregate Validation

Details:

The market segments could be based on household-level variables like auto-ownership categories, income categories or household size. Similarly, person-level segments can also be adopted using variables like age, gender, license status and employment status. Geographic variables such as area type also provide a useful categorization of the predicted data, which can then be compared to the observed statistics.

Example of Disaggregate Validation

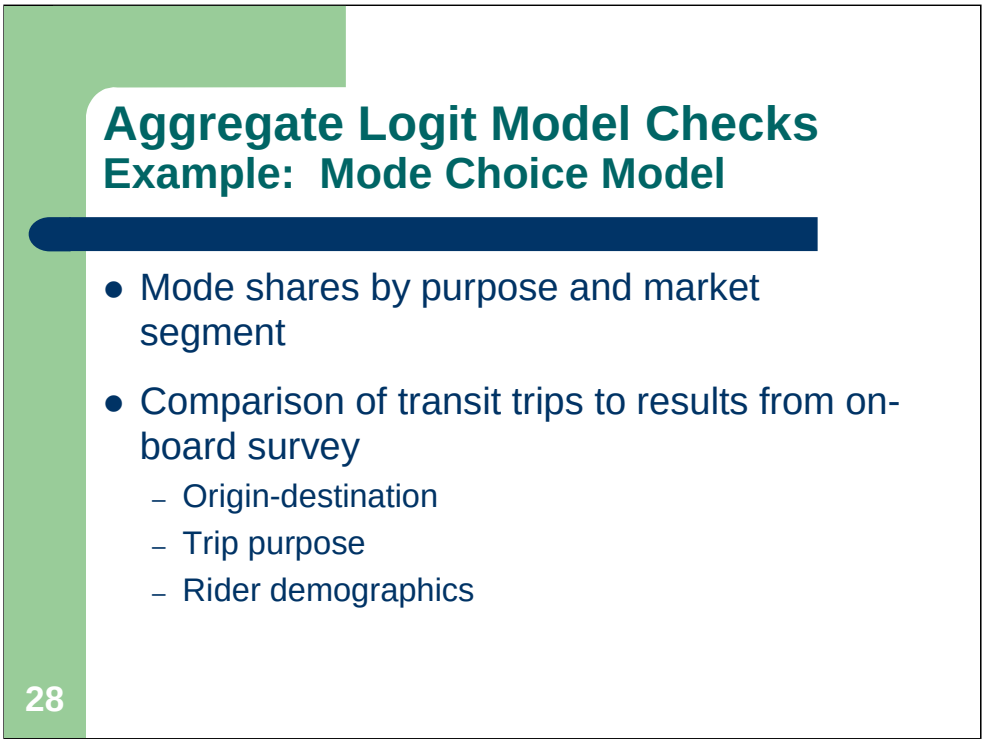
Choice	0	1	2	3+	Total
Non Motorized					
Number Chosen	47.4	104.5	270.2	158.5	580.6
Standard Deviation Chosen	7.5	14.1	26.5	19.0	36.3
	*A	V	A	A	A
Number Predicted	32.8	117.4	264.5	150.7	565.2
Auto Passenger					
Number Chosen	40.5	277.3	537.8	351.8	1,207.5
Standard Deviation Chosen	7.2	18.1	32.3	27.3	46.6
	V	***A	V	*V	A
Number Predicted	47.1	197.7	549.8	386.1	1,180.7
Drive Alone					
Number Chosen	0.0	1,265.9	4,225.5	3,233.4	8,724.8
Standard Deviation Chosen	0.0	25.4	44.8	35.7	62.7
		**V	A	A	A
Number Predicted	0.0	1,317.44	4,204.4	3,201.1	8,723.0
⋮	⋮	⋮	⋮	⋮	⋮
Total					
Number Chosen	119.3	1,770.7	5,326.2	3,928.7	11,144.9
Number Predicted	119.3	1,770.7	5,326.2	3,928.8	11,144.9

27

Key Contents: Disaggregate Validation – An Example

Details:

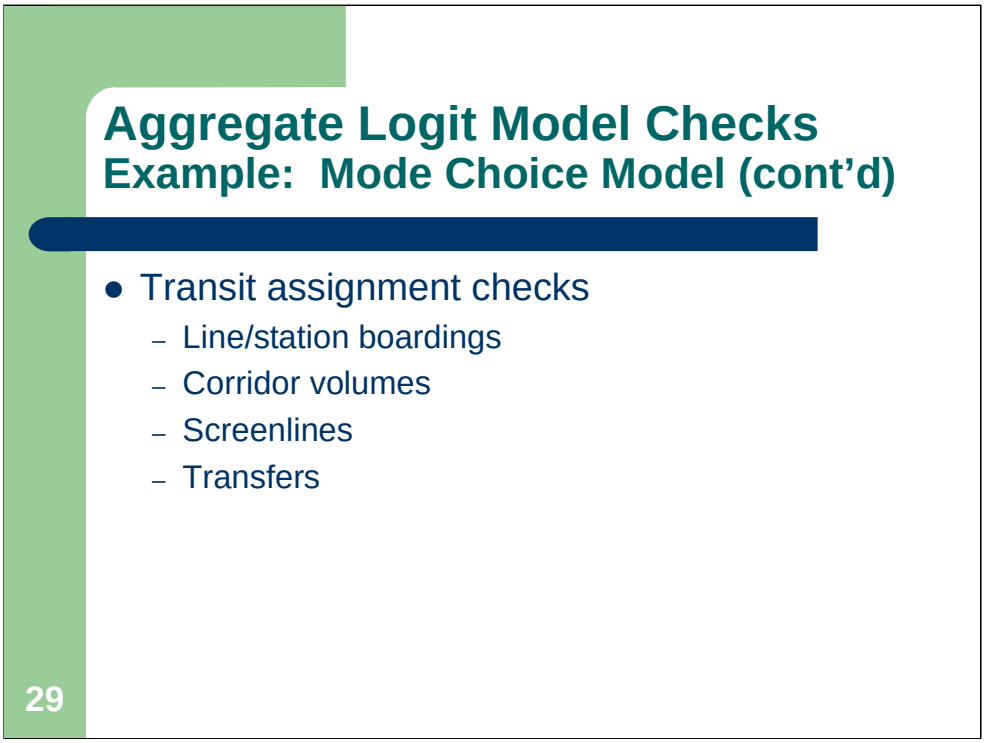
This slide provides an example of disaggregate validation using the estimation data. Here, the auto-ownership categories are used to tabulate the predicted proportion of modes, and these proportions are compared to the model's predictions.



Key Contents: Aggregate Validation

Details:

This is the most common check for the mode choice models. Here, we tabulate the predicted shares not at the individual level, but at the regional level. The mode shares are tabulated by origin-destination, trip purpose and rider demographics and these predicted shares are compared with data from other sources such as the household surveys and onboard surveys.

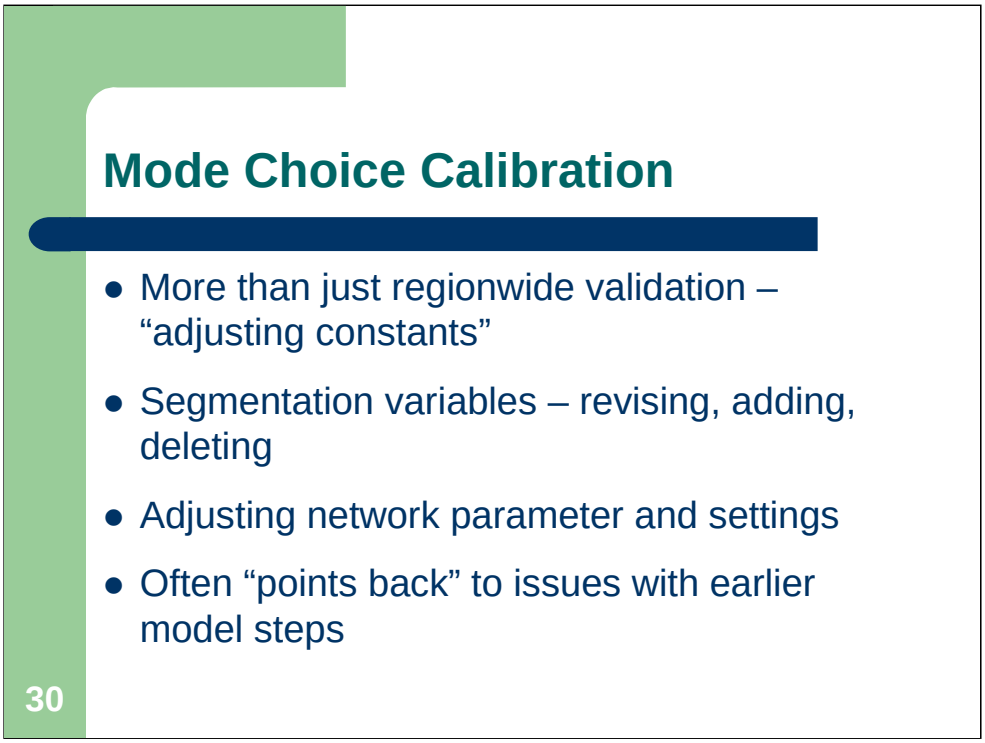


Key Contents: Aggregate Validation - Transit

Details:

For the transit portion, one could also check the post-assignment outputs such as the line or station-level boardings, and compare these numbers to those obtained from O-D surveys or even transit provider farebox counts. The checks could also be done at a screenline level.

Another key aspect of transit assignment checks is the transfer pattern. It is always a good idea to compare the modeled transfer behavior to the observed transfer behavior suggested by O-D surveys. Such comparisons yield useful insights into the necessary adjustment for mode choice parameters such as waiting time and transfer time coefficients.



Key Contents: Mode Choice Calibration

Details:

After the predicted mode shares are compared to the observed shares by trip purpose, the mode-specific constants can be adjusted to match the predicted and observed shares.

Of course, blindly adjusting the constants to match observed shares by market segment is not always advisable. As with any validation step, it is important to check the inputs first. The modeler needs to check the network skim settings and see if these are causing either high or low predicted shares.

These types of checks usually lead to necessary changes with earlier modeling steps such as the trip distribution step, for instance.

Vehicle Availabilty Model Validation Example

Variable	County							
	Yasasvi	Sarah	Gary	Tony	Chris	Sam	Lori	Region
% 0 Vehicles								
Observed	5%	13%	12%	7%	12%	7%	37%	17%
Model	3%	8%	8%	3%	5%	4%	24%	11%
% 1 Vehicles								
Observed	28%	35%	35%	31%	34%	33%	42%	35%
Model	27%	38%	38%	30%	34%	33%	49%	37%
% 2 Vehicles								
Observed	44%	37%	38%	44%	38%	43%	18%	34%
Model	47%	39%	39%	45%	43%	43%	21%	36%
% 3 Vehicles								
Observed	22%	14%	15%	18%	16%	17%	3%	13%
Model	22%	15%	15%	21%	18%	20%	5%	15%
Average Vehicles								
Observed	1.86	1.55	1.57	1.78	1.58	1.74	0.88	1.44
Model	1.95	1.65	1.65	1.90	1.79	1.82	1.11	1.59

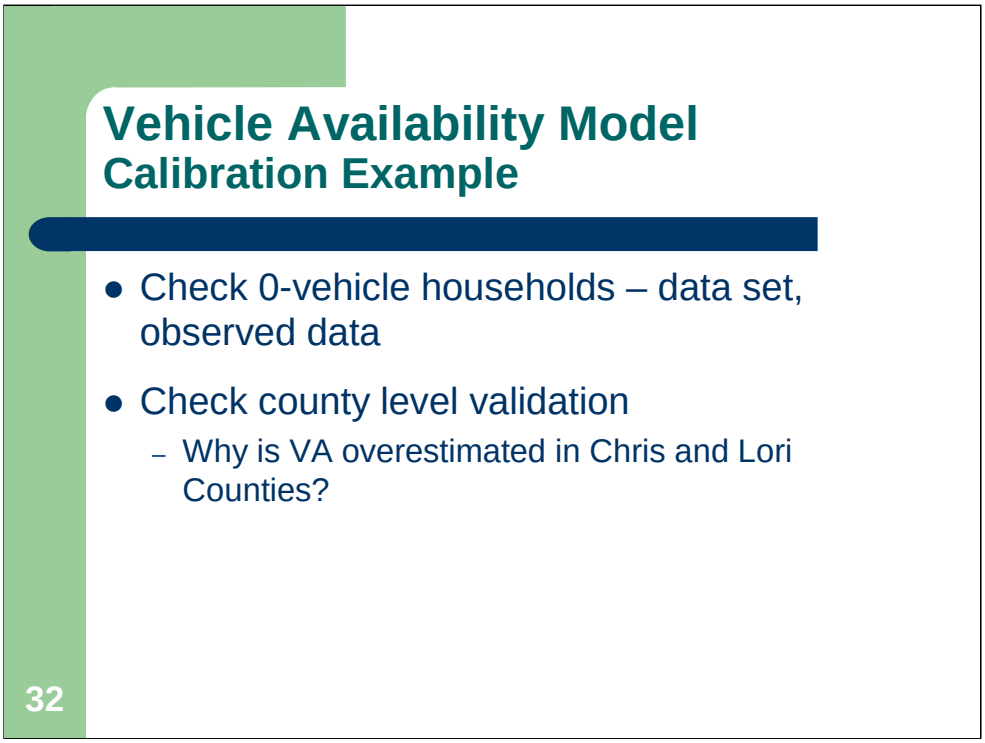
31

Key Contents: Vehicle Availability Model Validation

Details:

This slide shows the comparison of predicted and observed shares of 0-, 1-, 2- and 3+-vehicle household s by county.

One conclusion that can be drawn right away is that the share of 0-vehicle households are systematically underpredicted.

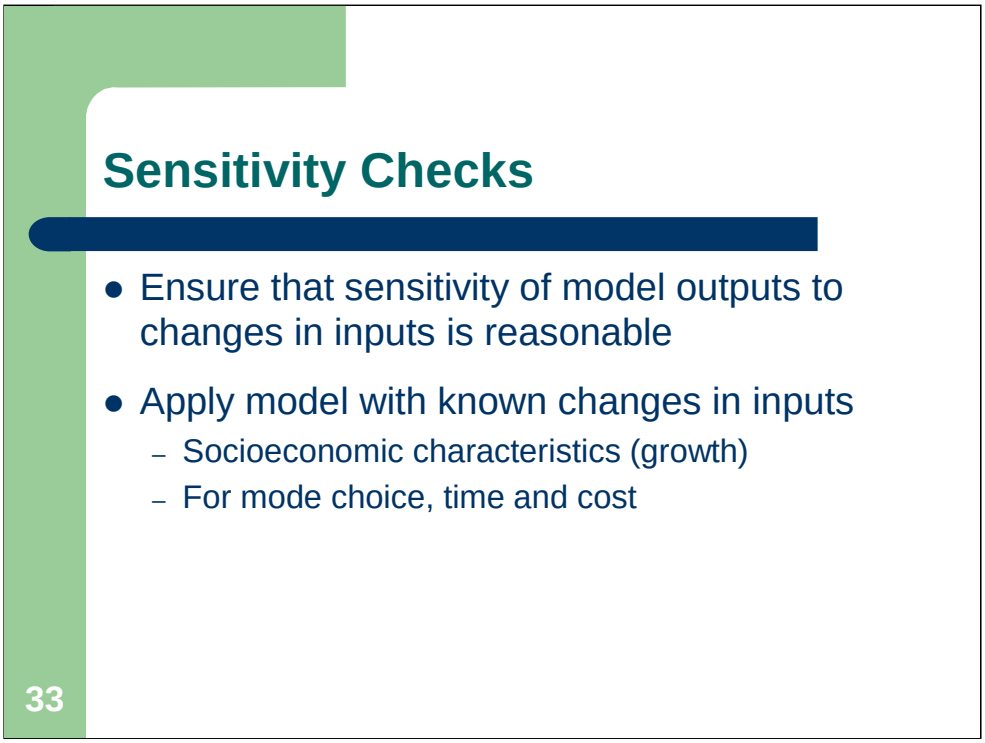


Key Contents: Vehicle Availability Model Calibration

Details:

Clearly, the results from the previous slides point to issues with the model for the 0-vehicle categories. A good starting point would be to look at Chris and Lori counties, where these differences are the most pronounced.

If there are no obvious data biases, one could adjust the constants for the 0-vehicle households to align the observed and predicted shares.



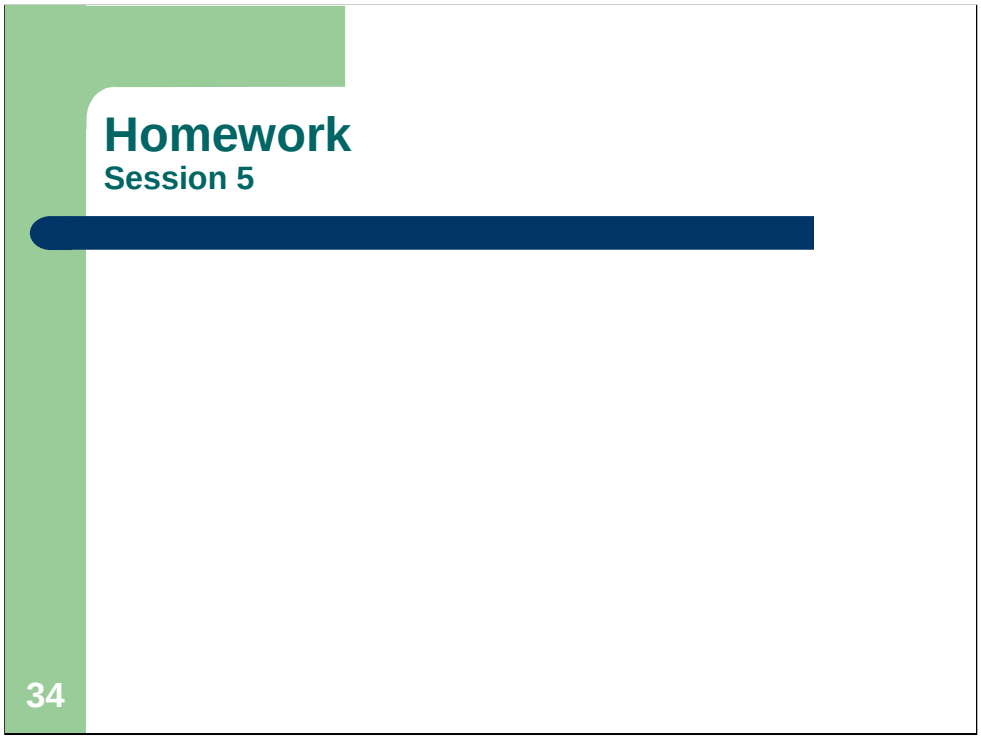
Key Contents: Sensitivity Checks

Details:

A key component of validation is sensitivity testing. So, the modeler should consider using various scenarios of socio-economic characteristics (aggressive growth, moderate growth, low growth), and apply the model to each of these scenarios. Using the results from each scenario, one can identify the sensitivity of the model to socio-economic characteristics.

Similarly, one could also test the sensitivity of various model components to assumptions about time and cost as well.

If the model appears over- or under-sensitive to the socio-economic or level-of-service inputs, the modeler will need to revisit the model and identify reasons and fixes.



Key Message: Homework

Details:

The homework for Session 5 can be downloaded from the course website. We would strongly recommend that the participants work through the homework problems to get more value out of this session.